

AD-A153 523

NONPARAMETRIC ESTIMATION OF THE PROBABILITY OF RUIN(U)  
WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER  
E W FREES FEB 85 MRC/TSR-2795 DARG29-80-C-0041

1/1

UNCLASSIFIED

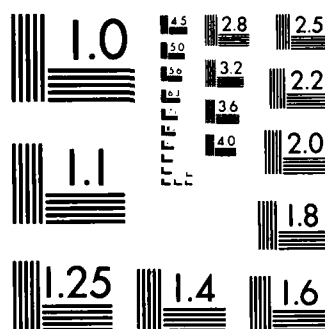
F/G 12/1

NL

END

FILED

DTIC



MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A153 523

MRC Technical Summary Report #2795

NONPARAMETRIC ESTIMATION OF  
THE PROBABILITY OF RUIN

Edward W. Frees

**Mathematics Research Center  
University of Wisconsin—Madison  
610 Walnut Street  
Madison, Wisconsin 53705**

February 1985

(Received January 30, 1985)

**DTIC FILE COPY**

Approved for public release  
Distribution unlimited

**DTIC  
ELECTE  
MAY 9 1985**

**S**

**D**

Sponsored by

U. S. Army Research Office  
P. O. Box 12211  
Research Triangle Park  
North Carolina 27709

85 4 10 120

UNIVERSITY OF WISCONSIN-MADISON  
MATHEMATICS RESEARCH CENTER

NONPARAMETRIC ESTIMATION OF THE PROBABILITY OF RUIN

Edward W. Frees

Technical Summary Report #2795  
February 1985

ABSTRACT

The finite and infinite horizon time probability of ruin are important parameters in the study of actuarial risk theory. This paper introduces procedures for directly estimating these key parameters from a random sample of observations without assumptions as to the parametric form of the distribution from which the observations arise. The estimators introduced apply to most of the classical models in which ruin probabilities are used and also apply to a much broader class of models. The procedures are based on the concept of sample reuse, an old idea in statistics which is becoming more popular due to the widespread availability of high speed computers. In this paper, the almost sure consistency of the estimators is established. Further, finite sample properties of the estimators are investigated in a simulation study.

AMS (MOS) Subject Classifications: Primary 62P05; Secondary 60G40, 62G99.

Key Words: reverse martingale, bootstrap.

Work Unit #4 (Statistics and Probability)

---

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

## SIGNIFICANCE AND EXPLANATION

The probability of ruin is the probability that claims against a risk-taking enterprise exceed its initial capital or reserve plus income at some point of time. If ruin occurs before a fixed time, such as one or ten years, this is called the finite horizon time problem and otherwise, the infinite horizon time problem. Even in the simplest models, ruin probabilities can be difficult to compute. There is a vast literature on this aspect of the problem dating back at least to the early 1900's.

In this paper it is shown how to approximate the probability of ruin directly from available claims data. Various properties of the approximation procedure are provided. Despite the long history of research in this area, it is not surprising that these procedures have not been developed previously because they are computer intensive and would have been computationally burdensome ten years ago. However, in today's era of fast computation, these procedures are inexpensive to implement. An example of the procedure is provided via simulating claims data and calculating approximations to the probability of ruin of the enterprise.

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/ _____	
Availability Codes	
Dist	Avail and/or Special
A-1	



The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

# NONPARAMETRIC ESTIMATION OF THE PROBABILITY OF RUIN

Edward W. Frees

## §1. INTRODUCTION

Let  $(X_i, Y_i)$   $i = 1, 2, \dots$  be i.i.d. random vectors with joint bivariate distribution function  $F$ . Use the random variable  $X_i$  to represent the  $i^{\text{th}}$  claim amount and the nonnegative random variable  $Y_i$  to represent the  $i^{\text{th}}$  interarrival time between claims. For  $t > 0$ , define the number of claims by time  $t$  as

$$N(t) = \sum_{k=1}^{\infty} I(Y_1 + \dots + Y_k \leq t)$$

where  $I(\cdot)$  is the indicator function. Premiums are assumed to arrive at a known steady rate, say,  $P$  per unit time. Thus, the amount that claims exceed premiums by time  $t$  is

$$U(t) = \sum_{k=1}^{N(t)} X_k - Pt. \quad (1.1)$$

Here we interpret the sum  $\sum_{k=1}^{N(t)}$  to be zero when  $N(t) = 0$ . We are primarily interested in the probability that  $U(t)$  exceeds an initial reserve  $u$  at some time  $t$  prior to or at  $T$ , the horizon time. This probability can also be defined by

$$\psi(u, T) = P\left(\sup_{0 \leq t \leq T} U(t) > u\right). \quad (1.2)$$

We are also interested in the probability that  $U(t)$  eventually exceeds an initial reserve  $u$ ,

$$\psi(u) = \psi(u, \infty) = \lim_{T \rightarrow \infty} \psi(u, T). \quad (1.3)$$

These probabilities,  $\psi(u, T)$  and  $\psi(u)$ , are called, respectively, the finite and infinite horizon time probability of ruin.

The probability of ruin is a key parameter in the collective theory of risk and has received considerable attention over the years. For several different types of

introduction to this theory, see Feller (1971), Beard et. al. (1984), Gerber (1979) and Bühlmann (1970). The behavior of the finite horizon time probability of ruin has been widely investigated not only because of its practical relevance but also because explicit computation of  $\psi(u,T)$  is difficult except in the most trivial cases. Explicit calculations have been given for some specific forms of  $F$ , cf., Seal (1978) and Thorin and Wikstad (1976). However, use of these explicit solutions has been limited due to their complexity and their dependence on a specific form of  $F$ . Because of these difficulties, papers giving approximations of  $\psi(u,T)$  suggested by limit theorems (e.g., as  $u \rightarrow \infty$ ,  $T \rightarrow \infty$ ) are abundant in the literature. The most successful of these approximations seem to be the diffusion approximations given by Siegmund (1979) and applied by Asmussen (1984). See Lalley (1984) for a refinement of Siegmund's work. Another class of methods for calculating  $\psi(u,T)$  is the Monte-Carlo method. Surprisingly, the Monte-Carlo method has received only limited attention in the risk theory literature. For some accounts, see Beard et. al. (1984) and Seal (1978). Note that this method does not depend on a specific form of  $F$  but does depend on complete knowledge of  $F$ .

The approximations of  $\psi(u,T)$  and  $\psi(u)$  given in this paper are different in nature from those sketched above and are inspired by the concept of sample reuse. Sample reuse is an old idea in statistics, popularized in the nineteen-forties by Hoeffding (1948) and more recently by Efron (cf., 1982). The idea for our applications is as follows. Consider the random variable,  $Z = \sup_{0 \leq t \leq T} U(t)$ . Since the distribution of  $U(t)$  is completely determined by  $F$ , then the distribution of  $Z$  is completely determined by  $F$ . Note that from (1.2),  $\psi(u,T) = 1 - P(Z \leq u)$ . By the usual multivariate Glivenko-Cantelli Theorem, knowledge of  $(X_i, Y_i)$   $i = 1, 2, \dots$  completely determines  $F$  and thus, for a sufficiently large sample size  $n$ , from  $\{(X_i, Y_i)\}_{i=1}^n$  we can build a good approximation to  $F$ , the usual (multivariate) empirical distribution function. This raises the natural question of how to build a reasonable estimator of  $\psi(u,T)$  based on a random sample of size  $n$ ,  $\{(X_i, Y_i)\}_{i=1}^n$ . The purposes of this paper are to argue that this is an important question and to develop estimators of  $\psi(u,T)$  (and  $\psi(u)$ ) that possess desirable properties.

In §2, estimators of  $\psi(u,T)$  and  $\psi(u)$  are introduced. The almost sure (a.s.) consistency of these estimators is proved in §3. A small simulation study is given in §4. Because of the nature of the approximations, the assumptions made in our development are different and more general than the usual risk theory assumptions. In §5 we discuss these differences and make other concluding remarks.



## §2. ESTIMATORS

Let  $\{(X_i, Y_i)\}_{i=1}^n$  be a random sample from a population with distribution function  $F$ . Let  $\{a_{1j}, a_{2j}, \dots, a_{nj}\} = A_j$  be the  $j^{\text{th}}$  permutation of  $\{1, 2, \dots, n\}$ ,  $j = 1, \dots, n!$ . We intend to reuse the sample by considering reordered pairs  $\{(X_{a_{ij}}, Y_{a_{ij}})\}_{i=1}^n$ ,  $j = 1, \dots, n!$

Define  $S_n = Y_1 + \dots + Y_n$ . For the  $j^{\text{th}}$  permutation, the number of claims by time  $t$  is

$$NA_j(t) = \sum_{k=1}^n I(Y_{a_{1j}} + \dots + Y_{a_{kj}} \leq t) \quad t \leq S_n$$

$$= n \quad t > S_n,$$

and the amount that claims exceeds premiums by time  $t$  is

$$UA_j(t) = \sum_{k=1}^{NA_j(t)} X_{a_{kj}} - Pt.$$

For the  $j^{\text{th}}$  permutation, we have ruin if

$$I(\sup UA_j(t) > u) = 1 \quad (2.1)$$

where the supremum is over the set  $\{t : 0 \leq t \leq \min(S_n, T)\}$ . Note that to compute the function in (2.1) one does not need to evaluate  $UA_j(t)$  at all  $t \in [0, \min(S_n, T)]$  but only at the random time points  $0, Y_{a_{1j}}, Y_{a_{1j}} + Y_{a_{2j}}, \dots, Y_{a_{1j}} + \dots + Y_{a_{nj}}$ . The first type of estimator we consider is the average over all permutations,

$$\psi_n(u, T) = (n!)^{-1} \sum_p I(\sup UA_j(t) > u) \quad (2.2)$$

and

$$\psi_n(u) = \psi_n(u, S_n) \quad (2.3)$$

where  $\sum_p$  is the sum over all permutations of  $\{1, 2, \dots, n\}$ . Note that we could alternatively define

$$\psi_n(u) = \psi_n(u, \infty) = \lim_{T \rightarrow \infty} \psi_n(u, T).$$

The consistency of these estimators is provided in the following

Theorem 2.1.

Suppose that  $EY > 0$ . Then, for each  $T$ ,

$$\lim_{n \rightarrow \infty} \psi_n(u, T) = \psi(u, T) \quad \text{a.s.} \quad (2.4)$$

and

$$\lim_{n \rightarrow \infty} \psi_n(u) = \psi(u) \quad \text{a.s.} \quad (2.5)$$

By the SLLN, the requirement  $EY > 0$  ensures that  $S_n \rightarrow \infty$  a.s. The proof, given in §3, is based on the idea that  $\psi_n$  can be shown to be a reverse martingale plus negligible terms. We remark that the estimators  $\psi_n$  defined in (2.2) and (2.3) each require the evaluation of  $n!$  indicators of ruin, an extensive amount of computations even for moderate sample sizes (say,  $n > 10$ ). Because of this computational difficulty, we now introduce a bootstrap estimator of the probability of ruin. The bootstrap methodology, a name coined by and a methodology popularized by Efron (cf., 1982), is also computer-intensive but does not require a prohibitive amount of computation.

Let  $B = B(n)$  be a positive integer depending on  $n$  such that  $B \rightarrow \infty$  as  $n \rightarrow \infty$ . Based on the observed sample  $\{(X_i, Y_i)\}_{i=1}^n$ , we draw  $B$  bootstrap realizations of  $\psi_n$  in the following two steps. For  $b = 1, \dots, B$ ,

Step 1. Make  $n$  independent (conditional on  $\{(X_i, Y_i)\}_{i=1}^n$ ) draws with replacement from  $\{(X_i, Y_i)\}_{i=1}^n$  to get  $(X_i^{*b}, Y_i^{*b})$  for  $i = 1, \dots, n$ .

Step 2. Define

$$N^{*b}(t) = \sum_{k=1}^n I(Y_1^{*b} + \dots + Y_k^{*b} \leq t)$$

and compute

$$\psi_n^{*b}(u, T) = I\left(\sup_{0 \leq t \leq \min(S_n, T)} \left(\sum_{k=1}^{N^{*b}(t)} X_k^{*b} - Pt\right) > u\right).$$

The bootstrap estimates are defined by

$$\psi_n^*(u, T) = B^{-1} \sum_{b=1}^B \psi_n^{*b}(u, T) \quad (2.6)$$

and

$$\psi_n^*(u) = \psi_n^*(u, S_n) . \quad (2.7)$$

The consistency of these estimators is provided in the following

Theorem 2.2.

Suppose  $EY > 0$  and

$$\log n = o(B^{1/2}) . \quad (2.8)$$

Then, for each  $T$ ,

$$\lim_{n \rightarrow \infty} \psi_n^*(u, T) = \psi(u, T) \quad \text{a.s.} \quad (2.9)$$

and

$$\lim_{n \rightarrow \infty} \psi_n^*(u) = \psi(u) \quad \text{a.s.} \quad (2.10)$$

The proof of Theorem 2.2 is given in §3. The condition on  $B$  in (2.8) guarantees that  $B$  grows sufficiently quickly to achieve a.s. convergence. In §4 we investigate the finite sample properties of  $\psi_n^*$  is a small simulation study.

### §3. PROOFS

Define  $Y_{1n}, Y_{2n}, \dots, Y_{nn}$  to be the order statistics of  $Y_1, Y_2, \dots, Y_n$  and let  $X_{in}$  be the claims amount associated with  $Y_{in}$ ,  $i = 1, \dots, n$ . Let  $G_n = \sigma((X_{in}, Y_{in}), i = 1, \dots, n, (X_i, Y_i), i > n)$  for  $n \geq 1$ , a nonincreasing sequence of sub  $\sigma$ -fields. Because all the arguments follow easily for the case  $T = \infty$ , we only give proofs for the a.s. consistency of  $\psi_n = \psi_n(u, T)$ . We now define a version of  $\psi_n$  and show that it is a reverse martingale. Later we show that this version is close to  $\psi_n$  in the appropriate sense.

For the  $j^{\text{th}}$  permutation of  $\{1, 2, \dots, n\}$ , let

$$NA_j^i(t) = NA_j(t) + \sum_{k > n} I(S_k \leq t)$$

be a version of the number of claims by time  $t$ ,  $j = 1, \dots, n!$ . With  $a_{kj} = k$  for  $k > n$ , define

$$\psi_n'(u, T) = (n!)^{-1} \sum_p I\left(\sup_{0 \leq t \leq T} \sum_{k=1}^{NA_j^i(t)} x_{a_{kj}} - Pt > u\right). \quad (3.1)$$

We have the following property for this version of the probability of ruin.

#### Lemma 3.1

For each  $u, T$ ,  $(\psi_n'(u, T), G_n)$  is a reverse martingale.

#### Proof:

It is easy to see that  $\psi_n' = \psi_n'(u, T)$  is  $G_n$ -measurable and  $G_n$ -integrable. With  $U(t)$  defined as in (1.1), we have the relation

$$\psi_n' = E\left\{I\left(\sup_{0 \leq t \leq T} U(t) > u\right) | G_n\right\}. \quad (3.2)$$

Thus,

$$\begin{aligned} E\{\psi_n' | G_{n+1}\} &= E\{E\{I(\sup_{0 \leq t \leq T} U(t) > u) | G_n\} | G_{n+1}\} \\ &= E\{I(\sup_{0 \leq t \leq T} U(t) > u) | G_{n+1}\} = \psi_{n+1}' \quad \dagger \end{aligned}$$

Proof of Theorem 2.1:

From (3.2), we have that  $E\psi'_n(u, T) = \psi(u, T)$ . From Lemma 3.1 and the (reverse) martingale convergence theorem, it is easy to show that

$$\lim_{n \rightarrow \infty} \psi'_n(u, T) = \psi(u, T) \quad \text{a.s.} \quad (3.3)$$

Define the stopping time

$$\tau = \inf \left\{ n \geq 1 : \sum_{k=1}^n (X_k - PY_k) > u \right\}.$$

Now, it is easy to see that  $\psi_n(u, T)$  and  $\psi'_n(u, T)$  differ only on the set  $\{n < \tau < \infty\}$ .

Thus

$$0 \leq \psi_n(u, T) - \psi'_n(u, T) = I(n < \tau < \infty) \rightarrow 0 \quad \text{a.s.}$$

This and (3.3) are sufficient for the proof.  $\dagger$

Proof of Theorem 2.2:

Denote  $\psi_n^* = \psi_n^*(u, T)$ . From (2.4), a sufficient condition for (2.9) is, for  $\epsilon > 0$ ,

$$\sum_n P(|\psi_n^* - \psi_n| > \epsilon) < \infty.$$

We show

$$\sum P(\psi_n^* - \psi_n > \epsilon) < \infty, \quad (3.4)$$

the proof of the other inequality being similar.

Define  $F_n = \sigma((X_i, Y_i), i = 1, \dots, n)$  for  $n \geq 1$ , a nondecreasing sequence of sub  $\sigma$ -fields. Now, from the condition (2.8), there exists a positive constant  $K$  such that, for sufficiently large  $n$ ,

$$n^{-1} > \exp(-KB^{1/2}). \quad (3.5)$$

From the Markov inequality, with  $\delta > \epsilon^{-1}$ , we have

$$P(\psi_n^* - \psi_n > \epsilon) \leq \exp\{-\epsilon\delta KB^{1/2}\} E\{\exp(\delta KB^{1/2}(\psi_n^* - \psi_n))\}.$$

From (3.5),  $\exp\{-\epsilon\delta KB^{1/2}\}$  is summable. Thus, to prove (3.4), we need only show

$$\sup_n E\{\exp(sB^{1/2}(\psi_n^* - \psi_n))\} < \infty \quad (3.6)$$

where  $s = \delta K$ . From (2.6), since conditionally on  $F_n$ ,  $\psi_n^*$  is the mean of a binomial random variable, we have

$$\begin{aligned} & E\{\exp(sB^{1/2}(\psi_n^* - \psi_n))\} \\ &= E\{(1 - \psi_n)\exp(-s\psi_n B^{-1/2}) + \psi_n \exp(s(1 - \psi_n)B^{-1/2})\}^B \\ &= E(1 + s^2\psi_n(1 - \psi_n)/(2B) + o(B^{-3/2}))^B \\ &< (1 + s^2/(8B) + o(B^{-3/2}))^B \end{aligned}$$

by a Taylor-series expansion. This is sufficient for (3.6) and hence (2.9). †

#### §4. SIMULATION

In this section, finite sample properties of the bootstrap estimators introduced in §2 were investigated. A simple example was used so that calculation of exact probabilities of ruin and comparison with other studies were possible. Claim amounts were assumed to be exponentially distributed with mean 1. The claims were assumed to arrive as a Poisson process with intensity  $\rho = .8$ , i.e., interarrival times are independent and exponentially distributed with mean 1.25. The claims amount and arrival times were assumed to be independent and premiums arrive with unit intensity ( $P = 1$ ). Under these assumptions, it can be shown that

$$\psi(u) = .8 \exp\{-.2u\}, \quad (4.1)$$

thus giving an easy expression for exact values of  $\psi(u)$ . Furthermore, in a recent study, Asmussen (1984) has provided exact values of  $\psi(u, T)$  and several popular approximations of  $\psi(u, T)$  for various values of  $T$ .

The bias (BIAS) and root mean square error (RMSE) were used to judge the performance of the estimators. All computations were done on a VAX 11/750 owned and operated by the Department of Statistics at the University of Wisconsin-Madison. The IMSL Fortran subroutines produced the random deviates.

In Table 1 we give the results of the performance of the bootstrap estimator of the infinite horizon time probability of ruin. The tables give the criteria for ruin probabilities  $\psi(u) = 1\%, 5\%, 10\%, 40\%$  and for sample sizes  $n = 10, 30, 50, 100, 150$ . The ruin probabilities were chosen to represent a range which is typically of interest to the actuarial community. The sample sizes were selected to represent small and moderate numbers of claims. For small sample sizes  $n = 10, 30, 50$ ,  $B = 200$  bootstrap replications were run to compute the estimators for each simulation trial. For larger sample sizes  $n = 100, 150$ , we used  $B = 100$  bootstrap replications to compute the estimators. Using a smaller  $B$  produced considerable savings in computer run time and did not seem to affect significantly the accuracy of the results.

In Table 2 we give the results of the performance of the bootstrap estimator of the finite horizon time probability of ruin. The number of bootstrap replications are as in

Table 1. To make our study comparable to the study of Asmussen (1984), the level of initial reserve yielding  $\psi(u) = .8\%$  and horizon times  $T = 13.8, 41.3, 68.8, 96.4$  were selected. For this level of reserve and these times  $T$ , the exact probabilities  $\psi(u, T)$  were taken from Asmussen (1984). Because  $\psi(u)$  is small, we used larger sample sizes,  $n = 30, 50, 100, 150, 200$ .

TABLE 1 - BOOTSTRAP ESTIMATOR OF  $\psi(u)$

$\psi(u)$	n	B	BIAS	RMSE
1%	10	200	-.00982	.00999
			-.00773	.01770
			-.00364	.02116
	50	100	.00470	.03910
	100		.01365	.05615
	150			
5%	10	200	-.04718	.04861
			-.02602	.06063
			-.01025	.07711
	50	100	.01065	.10527
	100		.02400	.11267
	150			
10%	10	200	-.09023	.09648
			-.03965	.10944
			-.01650	.12108
	50	100	.00960	.14964
	100		.02000	.14489
	150			
40%	10	200	-.24840	.32200
			-.08825	.26922
			-.04225	.24297
	50	100	-.04010	.22975
	100		-.02250	.21270
	150			



TABLE 2 - BOOTSTRAP ESTIMATOR OF  $\psi(u, T)$

T	$\psi(u, T)$	n	B	BIAS	RMSE
13.8	.00007	30	200	.00118	.00546
		50		.00138	.00459
		100	100	.00043	.00346
		150		.00048	.00272
		200		.00023	.00173
41.3	.00145	30	200	.00005	.00581
		50		.00330	.01590
		100	100	.00805	.02699
		150		.00670	.02276
		200		.00370	.01144
68.8	.00338	30	200	-.00188	.00611
		50		.00140	.01566
		100	100	.00892	.03525
		150		.01387	.04527
		200		.01017	.03118
96.4	.00491	30	200	-.00341	.00674
		50		.00013	.01560
		100	100	.00739	.03489
		150		.01509	.05172
		200		.01579	.04764

As was expected, the performance of the infinite time estimator was better the closer  $\psi(u)$  was to 50%. In Table 1, the asymptotic theory comes quickly into play when  $\psi(u) = 40\%$  as evidenced by the decreasing BIAS and RMSE terms with increasing sample size. For  $\psi(u) = 10\%$ , we see some levelling off of the RMSE term from  $n = 100$  to  $n = 150$ . The smaller probabilities showed no evidence of levelling off for the sample sizes considered. Although the motivation of decreasing B at higher sample sizes was to

decrease computing costs, this did not seem to affect the performance of the estimators. See Efron (1982) for further discussion of the selection of  $B$ .

Perhaps the most interesting fact of the simulation study was that the performance of the finite time estimators improves as  $T$  decreases. In Table 2, the asymptotic theory comes quickly into play when  $T = 13.8$ . For  $T = 41.3$ , we see the decreasing trend in the BIAS and RMSE terms beginning at a larger sample size  $n = 100$ . The intuition is as follows. For  $T = 13.8$ , on the average it requires about 11 observations to check for ruin by time 13.8 (since  $11 \times 1.25 \times 1 = 13.75$ ). With a sample size of  $n = 100$ , we have approximately 9 independent and identical realizations of an indicator of ruin by time  $T = 13.8$ . Repeating this reasoning for  $T = 41.3$ , the reader can verify that we have only approximately 3 independent and identical realizations of this indicator of ruin. While the goal of the estimators introduced in §2 is to reuse sampling information, by increasing the horizon time  $T$  we increase the number of observations necessary to check for ruin and thus reduce the number available for resampling.

## §5. DISCUSSION

The estimators introduced in §2 are defined in terms of the classical method of collecting data, i.e., based on a random sample of size  $n$ . However, it is easy to modify these estimators for other methods of data collection and retain their important statistical properties. For example, suppose an insurance company would like to estimate the probability of ruin based on one year of observed data. Then the sample size itself is random. However, it is easy to see that the usual results on random change of time (cf., Csörgö and Révész, 1980, Theorem 7.1.1) can be applied to preserve the properties stated in Theorems 2.1 and 2.2. In the risk and queuing literature, this change of time is usually referred to as a transition to operational time.

The most important drawback of the estimators  $\psi_n$  and  $\psi_n^*$  is that they rely heavily on the independence of the bivariate pairs  $(X_i, Y_i)$ . While this assumption is used in most models constructed to calculate the probability of ruin, other models such as a model which

uses the mixed Poisson process for the claims number process (cf., Seal and Gerber, 1984) do not.

Except for the important assumption of independence, the estimators  $\psi_n$  and  $\psi_n^*$  and their properties are relatively free from assumptions when compared to other estimators of the probability of ruin. Both estimators are nonparametric in the sense that they do not assume a particular parametric form nor knowledge of the distribution function  $F$ . In the assumptions we have not precluded the case  $P(Y_i = 0) > 0$ , thus allowing for the possibility of multiple claims at any point in time. Further, we have not assumed independence between the claim arrival time and the claim amount. If this assumption is made, we conjecture that a different estimation procedure can be constructed that uses the data more efficiently, in some sense. We leave this as an open question for future research.

We remark here that questions of finding the optimal or most efficient procedure, in some sense, have not been addressed. For example, in Step 1 of the definition of  $\psi_n^*$  in §2, if the drawing of  $\{(X_i^{*b}, Y_i^{*b})\}_{i=1}^n$  was made without replacement then inspection of the proof of Theorem 2.2 shows that Theorem 2.2 still holds. Whether the drawing should be made with or without replacement we leave as an area of future research. The bootstrap estimator  $\psi_n^*$  is computationally similar to the usual Monte-Carlo procedure. However, it is different philosophically in that the Monte-Carlo procedure assumes knowledge of the underlying distribution function while the bootstrap does not. This similarity suggests that dynamic factors in ruin probabilities such as interest, inflation, economic cycles, etc., that have been incorporated in Monte-Carlo methods (cf., Beard et. al., 1984, Chapter 7) may be incorporated in bootstrap estimators. We leave this as an open area for future research.

Acknowledgement. The author would like to thank Professors James C. Hickman and Robert B. Miller for helpful comments on an earlier draft of this paper.

# REFERENCES

- [1] Asmussen, S. (1984). Approximations for the probability of ruin within finite time. Scandinavian Actuarial Journal 31-57.
- [2] Beard, R., Pentikäinen, T. and Pesonen, E. (1984). Risk Theory: The Stochastic Basis of Insurance. Chapman and Hall, New York.
- [3] Bühlmann, H. (1970). Mathematical Methods in Risk Theory. Springer-Verlag, Heidelberg.
- [4] Csörgö, M. and Révész, P. (1981). Strong Approximations in Probability and Statistics. Academic Press, New York.
- [5] Efron, B. (1982). The Jackknife, the Bootstrap and Other Resampling Plans. SIAM, Philadelphia.
- [6] Feller, W. (1971). An Introduction to Probability Theory and Its Applications, Vol. 2. Wiley, New York.
- [7] Gerber, H. (1979). An Introduction to Mathematical Risk Theory. Irwin, Homewood, Illinois.
- [8] Hoeffding, W. (1948). A class of statistics with asymptotically normal distributions. Annals of Mathematical Statistics 19, 293-325.
- [9] Lalley, S. (1984). Limit theorems for first-passage times in linear and nonlinear renewal theory. Advances in Applied Probability 16, 766-803.
- [10] Seal, H. (1978). Survival Probabilities. Wiley, New York.
- [11] Seal, H. and Gerber, H. (1984). Mixed Poisson processes and the probability of ruin. Insurance: Mathematics and Economics 3, 189-190.
- [12] Siegmund, D. (1979). Corrected diffusion approximations in certain random walk problems. Advances in Applied Probability 11, 701-719.
- [13] Thorin, O. and Wikstad, N. (1976). Calculation of ruin probabilities when the claim distribution is lognormal. Astin Bulletin 9, 231-246.

EWf/ed

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 2795	2. GOVT ACCESSION NO. <b>AD-A153523</b>	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle)  NONPARAMETRIC ESTIMATION OF THE PROBABILITY OF RUIN		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s)  Edward W. Frees		8. CONTRACT OR GRANT NUMBER(s)  DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 4 - Statistics and Probability
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P.O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE February 1985
		13. NUMBER OF PAGES 15
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report)  UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report)  Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)  reverse martingale, bootstrap		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)  The finite and infinite horizon time probability of ruin are important parameters in the study of actuarial risk theory. This paper introduces procedures for directly estimating these key parameters from a random sample of observations without assumptions as to the parametric form of the distribution from which the observations arise. The estimators introduced (cont.)		

ABSTRACT (cont.)

apply to most of the classical models in which ruin probabilities are used and also apply to a much broader class of models. The procedures are based on the concept of sample reuse, an old idea in statistics which is becoming more popular due to the widespread availability of high speed computers. In this paper, the almost sure consistency of the estimators is established. Further, finite sample properties of the estimators are investigated in a simulation study.

**END**

**FILMED**

**6-85**

**DTIC**